3/PRTS

09/786366

JC08 Rec'd PCT/PTO 15 MAR 2001

METHOD AND APPARATUS FOR VIDEO SIGNAL PROCESSING

The present invention relates to a method for video processing, wherein possible movement between successive fields of the images split into even and odd is detected and the mode, i.e. video mode or film mode is determined.

More in particular, the invention relates to the processing of video signals from a video tape, a DVD, a laser disc, a tuner, etc. and therefore of standard PAL, SECAM and NTSC video signals.

Given that, in particular in the case of film mode, be it so-called "2:2 pull-down" or so-called "3:2 pull-down", which involve the film images being split into even and odd fields, the quality of the image on the monitor can exhibit annoying effects if these signals are processed in the classical manner, it is known to determine, via movement detection, between successive fields, according to what mode the current video signal was converted in order then to carry out specific further processing of the film mode signals.

With known methods, the movement detection is carried out by detecting possible movement of the edges of the fields, the result being that rapid intervention will take place in the event of many details in a field. This requires a large number of auxiliary devices such as a battery of correlators.

Employing these known methods therefore requires a relatively expensive and bulky apparatus.

It is an object of the invention to provide a method for video processing which avoids these and other drawbacks and which, even in the case of film mode, is able to display a high-quality image on the monitor,

while using a relatively inexpensive apparatus which can be of compact design.

This object is achieved according to the invention by the sequence of movement or standstill between successive fields being detected and this sequence over a number of fields being stored in a memory, followed by said sequence being compared with patterns inherent to the mode and, if ordinary video mode is detected, median filtering is carried out whereas, if film mode (2:2 pull-down or 3:2 pull-down) is detected, the median filtering is switched off and, in synchronization with the film phase, the even and odd fields which match and are derived from one and the same film image are merged again until the original film image is obtained and said image is repeated until again a following original film image can be constructed by means of the abovementioned merging.

The movement detection, i.e. the detection of the abovementioned sequence of movement and standstill, can be performed by a three-point median filtering operation, after which the result of said median filtering operation and the incoming information of a following field is filtered by two low-pass filters, the absolute difference of the result of these two low-pass filters is calculated and the differences are summed, the sum, possibly divided by a number, being compared with a threshold value, the result of said comparison forming the abovementioned sequence, stored in a memory, of the movement and standstill between successive fields.

The video signals to be processed are preferably subjected to doubling, and preferably also quadrupling and preferably also field rate doubling.

The invention also relates to an apparatus which is particularly suitable for employing the method according to any one of the above embodiments.

The invention therefore relates to an apparatus which includes a movement detector, a film mode/video mode detector connected thereto, synchronization means to synchronize processing with the film phase, and a film processor proper.

Preferably, the movement detector is connected to a median filter having as inputs the current field and the next field of the video images, and includes two low-pass filters, one of which connects to the output of the median filter and the other has as an input the information of the next incoming field, a differentiator which connects to the two low-pass filters to calculate the difference between the outputs of these, a summator connected to said differentiator, a counter connected to said summator and a comparator connected thereto for comparing the output of the counter with a threshold value.

Preferably, the film mode/video mode detector includes a shift register in which the result of the comparator over a number of fields is stored, so that a comparison is possible with a pattern inherent to a specific mode.

The apparatus may include means for employing doubling and/or means for performing quadrupling, and a field rate converter from 50 or 60 Hz to 100 or 120 Hz.

With a view to providing a better demonstration of the characteristics of the invention, preferred embodiments of a method and apparatus for video processing according to the invention are described hereinafter, as an entirely non-limiting example, with reference to the accompanying figures in which:

Figure 1 represents a block diagram of an apparatus used for implementing the method according to the invention;

Figure 2 schematically represents the 2:2 pull-down conversion and processing;

Figure 3 schematically represents the movement sequence between the fields in the case of film processed by 2:2 pull-down;

Figure 4 schematically represents the 3:2 pull-down conversion and processing;

Figure 5 schematically represents the movement sequence between the fields in the case of film processed of 3:2 pull-down.

Starting with interlaced video having 525 lines at 60 Hz or 625 lines at 50 Hz, the interlacing is firstly moved and the video is therefore first converted to non-interlaced video.

For this purpose use is made of a known or progressive scanning technique, in which a three-point median, obtained with the aid of median filter 1, carries out the compensation for movement.

In the case of interlaced video, each frame consists of two fields, successive even and odd fields being shifted by half a line with respect to one another. The three-point median has as its input the pixels lying vertically above one another of two successive lines from the current field coming in via 1A and the pixel vertically below of the intermediate line of the following field coming in via 1B.

The result of the median processing, combined with the original lines of the current image, provides a movement-compensated image of 625 or 525 lines, non-interlaced, and 50/60 Hz.

The result of the median filter 1 is stored in a memory bank (display bank) 2 which consists of two image FIFOs 3 and 4 which are written in interlaced form. The even pixels go to FIFO 3, the odd pixels to FIFO 4.

"Doubling" is the constant read-out of the display bank 2 at the normal rate.

After interlacing has been removed, a stable image without line flicker is obtained. To obtain optimal screen coverage on CRT-based monitor systems, a quadrupling technique is used additionally.

Simultaneously with the calculation of the median in the median filter 1, vertically intermediate pixels are calculated via vertical linear interpolation in the interpolator 5, which connects, via 5A to the output of the median filter 1. Again, 625/525 lines of non-interlaced 50/60 Hz are obtained.

The result of this interpolation is stored in a memory bank (display bank) 6 which, like the abovementioned, consists of two image FIFOs 7 and 8 which are written in interleaved form. The even pixels go to FIFO 7, the odd pixels to FIFO 8.

To obtain a quadrupled image, a line of the display bank 2 and a line of the display bank 6 is read alternately by a memory controller 9, the result being relayed to a digital-to-analogue converter 10.

Each line consists of 1024 pixels, 512 even pixels from FIFO 3 or 7, and 512 odd pixels from FIFO 4 or 8, which are read out in interleaved form and in the correct pixel phase. Ultimately, a total image of 1250 or 1050 lines, non-interlaced, is obtained a line frequency of 64 kHz and a frame rate of 50 or 60 Hz, respectively.

To render the plane flicker inherent to low frame rates of 50 or 60 Hz invisible, recourse is moreover had to field rate upconversion, which involves increasing the frequency to 100 to 120 Hz.

In a first embodiment, the lines of the abovementioned display bank 2 are read out by the memory controller 9 twice in succession and at twice the normal rate and therefore at a frequency of 100 or 120 Hz instead of 50 or 60 Hz, an image being formed which has 625 or 525 lines, respectively, non-interlaced, displayed at double the frame rate. The line frequency is then 64 kHz.

To retain the advantage of better screen filling, it would be necessary to change to a line frequency of 128 kHz, which cannot be achieved by most display equipment. To solve this problem, the memory controller 9, in a second embodiment, first reads out the entire contents of the display bank 2 at twice the normal rate and then, at twice the normal rate, reads out the entire contents of the display bank 6. Since both display banks 2 and 6 are alternately read out at twice the rate, a resulting field rate of 100 or 120 Hz will again be obtained.

Since the contents of the display bank 2 differ spatially from the content of the display bank 6, it is necessary to ensure that the data of both display banks 2 and 6, which do agree temporally, are not written to the same position on the monitor. To achieve correct interlacing at 100 or 120 Hz, the outgoing vertical frame pulse (100/120 Hz) is then, every other frame, shifted upwards by half a 64 kHz line.

As a result of the interlacing, the information from the one display bank ends up in between the information of the other display bank 6, which produces optimal

screen coverage within the same time as with normal quadrupling (20 ms or 16.6 ms).

This type of image, because of the interlacing using non-equidistant frame pulses instead of equidistant frame pulses, can be used only on equipment which supports this type of interlacing, such as HDTV-compatible equipment.

The frame pulse for the second 100 or 120 Hz field must be calculated. The length of the frame is measured on the basis of the number of 64 kHz lines in an original field. The second frame pulse is positioned in the middle between two original frame pulses. Without interlacing, equidistant 100 or 120 Hz frame pulses will be obtained.

The field rate upconversion can therefore be applied both to the doubling technique and to the quadrupling technique, with the option of doubling and quadrupling being employed separately, in conjunction with field rate upconversion or without it. In the case of doubling, the display bank 2 is constantly read out at normal rate, in the case of quadrupling every other line of the display banks 2 and 6 are read out at normal rate, in the case of doubling and field rate doubling the display bank 2 is read out constantly at double the rate, whereas in the case of quadrupling and field rate doubling, complete fields are read out alternately from the display banks 2 and 6 at double the rate, followed by interlacing with respect to the frame pulse.

Apart from video mode, i.e. normal video recordings in which movement can take place between each field (20 or 16.6 ms interval), the apparatus is able to process transmissions in other modes, including film mode or derivatives thereof such as video editing mode transmissions, in which mixing takes place of normal

video images with film mode images, and film editing mode, in which the input consists entirely of film mode images, but the correct sequence is missing.

Film mode can take the form of so-called "2:2 pull-down", in which 24 film images per second are converted into PAL 50 Hz with 50 images per second or 30 film images per second are converted into NTSC 60 Hz with 60 images per second, or of so-called "3:2" pull-down", where 24 film images per second are converted into NTSC 60 Hz and therefore 60 images per second.

In the case of "2:2 pull-down", the film, in order to be transmitted in the PAL transmission standard, is slightly accelerated to 25 images per second, and each film image or frame is then split into an even and an odd image or field, so that two fields A and B or C and D or E and F etc. are obtained which match one another and occupy the same temporal position but are spatially situated on a different location, i.e. are separated, as seen vertically, by half a scan line, as represented in the top section of Figure 2.

Each field A, B, C etc. consists of lines, 312.5 lines in the case of PAL, which in Figure 2 exist as A_0, A_1, A_2, etc. for field A, B_0, B_1, etc. for field B, etc. The original film image can be restored by the even and odd field successively being shown and interlaced on the monitor.

Likewise, to be able to transmit a film having 30 images per second in NTSC standard, each image or frame is analogously split into an even and an odd field.

In the case of "3:2 pull-down", as represented in the top section of Figure 4, frames consisting of an even and an odd field, for example field A and field B, are composed from, on the one hand, even and odd fields from the same instant and, on the other hand, an even

field of one instant and an odd field of the other instant, for example from odd field A, even field A', odd field A, even field B, odd field B', etc, so that ultimately 60 images per second are again achieved.

Starting from six original image frames, 15 interlaced fields are produced. Successively, three images or frames are split into an even and odd field obtained, followed by two split frames comprising an even field of one image and an odd field from another image, the upshot being that for each five successive fields, one field is repeated.

Film transmissions and therefore 2:2 and 3:2 pull-down images will, given correct detection and processing, give qualitatively much better results that normal video images.

In the above-described de-interlacing, use is made of a median filter 1 to obtain correct movement compensation. A median filter has the drawback, however, that it cannot readily cope with lines moving diagonally with the result that small steps or serrations, so-called "jaggies", are produced on the diagonals.

This is a highly annoying effect, which is even more noticeable in film transmissions in 2:2 or 3:2 pull-down form.

To ensure that this effect is markedly reduced with film transmissions, the de-interlacing must be effected without a median filter.

In the first instance, therefore, a movement detector 11 is used to detect between which fields there is movement and between which fields there is absolutely no movement.

This is done by observing correlations between these fields.

To this end prefiltering is performed as a first step.

Successive fields A (comprising lines A_0, A_1, etc) and B (comprising lines B_0, B_1, etc), which may match, do come from the same instant but, as a result of the image having been split vertically when film was converted to video transmission, they differ spatially.

Consequently it would be wrong for the pixel X from line A_0 to be compared with pixel X from line B_0, given that these are spatially separated; instead, the result of the abovementioned median filter 1, i.e. the median of pixel X of A_0, A_1 and B_0 will be compared with the pixel X from line B_0, since this median spatially coincides with pixel X from line B_0.

This is enough to obtain a first improvement in being able to find possible correlation between a field A and a field B.

The result of the median filter 1, on the one hand, and the information, coming in via 12, from the next field, on the other hand, are each filtered by means of a low-pass filter 13 or 14, for example a relatively inexpensive recursive filter, which continuously averages a number of successive pixels, for example 32, in accordance with the equation:
$(32.x_{old} - 1.x_{old} + 1.x_{new})(1/32)$.

The movement correlation is performed on the low frequencies and not on sharp transitions or details, in order not to be too strongly dependent on the net bandwidth of the input signal and on the possible errors of the median filter 1.

Then, for each pixel, the absolute difference is calculated, in the differentiator 15, between the result of the low-pass filter 13 for the result of the median filter, and the low-pass filter 14 for the incoming next image.

The successive absolute differences are summed in the summator 16, and each time there is an overflow of, for example, 2^9 or 512, the counter 17 connected thereto is incremented. What this amounts to is that the total sum of the absolute differences is divided, for example, by 512 in order to limit the magnitude of the final number.

This process is carried out over an entire field, but preferably also limited or keyed by so-called "windowing" from a specific start line to a specific stop line and from a specific start pixel to a specific stop pixel.

At the end of each field, the output of the counter 17 yields a number which, in principle, indicates that degree of movement between two successive video fields.

In a third step a relevant criterion is defined for determining whether there is indeed movement or standstill.

In the event of movement of bright images, the result of the sum of the absolute differences between two successive fields will, on average, be higher than in the case of dark fields. Consequently, the calculation of the threshold value for the movement criterion takes account of the luminance value of the current image being processed.

At the same time as the sum of the absolute differences, therefore, the total sum of the luminances is also calculated. The threshold value used for

movement is an empirically determined fraction of the sum of the luminance values, for example 1/64th. Depending on the current processing operation, this threshold value can be doubled or halved.

If there are small details in the one image or field, which shift with respect to the next image, for example noise or detailed structures, this can have a disruptive effect on the result of the movement detection.

If the threshold value comes out too low, the result, even with instantaneously standstill between successive images or fields, may nevertheless be movement, this being incorrect. Because of this, that information is filtered out by means of the abovementioned low-pass filters 13 and 14 in the prefiltering operation.

In as much as an edge boost, to be described hereinafter, is employed prior to movement detection taking place, which may be required for architectural reasons in the case of quadrupling, details of fine structures may after all play a significant part in determining the threshold value so that it is necessary, in calculating this threshold value, to take account of the status of the edge boosting device. The more pronounced the edge boost, the higher will the ultimate threshold value be.

This additional adjustment of the threshold value as a compensation for the edge boost is carried out simultaneously on the nominal value, for example 1/64th of the luminance value, double the value and half the value.

The threshold value is then tailored to the processing mode.

If this processing mode is the detection mode with continuous monitoring of the incoming data, the movement detection is set to nominal sensitivity, for example 1/64th of the luminance value.

Should some film mode or other have been detected and the processing have been tailored thereto, the movement detection is set to low sensitivity, for example twice the nominal threshold value.

In the case of a third mode, the film resynchronization mode (referred to as "film resync mode"), where the system is still in film mode but the movement detection in a single instance has detected an incorrect film phase, the processing reverts to standard median processing and the movement detection is set to high sensitivity, for example half the nominal threshold value, to enable rapid detection.

The abovementioned threshold value for movement is determined at the end of each field.

The film mode/video detection proper then takes place as follows:

When the abovementioned result of the counter 17 exceeds said threshold value, there is movement between the successive fields, and comparator 18, connected to the counter 17, of the movement detector 11 will generate a logical "1" which is clocked into a shift register 19.

When, conversely, the abovementioned sum drops below the threshold value, there is correlation or no movement and the comparator 18 of movement detector 11 will detect a logical "0" which is likewise entered into said shift register 19.

The shift register 19 has a length of a number of bits, for example eleven bits, which means that the history of the movement detector 11 can be measured over a number of fields, for example eleven fields.

In the case of the film mode/video detection, a detector 20 scans the sequence of the shift register 19 for a specific pattern inherent to the various modes.

The output of the detector 20 is connected, via 20A to the comparator 18 for feedback of the threshold value which, in fact, is determined in said detector 20.

In the event of movement in the case of film processed by 2:2 pull-down, i.e. where a film containing 24 images per second was converted to 50 images per second in PAL and transmission standard 50 Hz or a film with 30 images per second was converted to 60 images per second in NTSC 60 Hz, the movement sequence will be 0101010101..., as represented in Figure 3.

When the detector 20 therefore observes a 01010101.. sequence, the decision will be that a "2:2 pull-down mode" obtains.

In the event of movement in the case of film processed by 3:2 pull-down, i.e. where a film containing 24 images per second was converted to 60 images per second in NTSC 60 Hz, the movement sequence will be 1010010100101..., as represented in Figure 5, so that, when the detector 20 therefore observes this sequence, the decision will be that a "3:2 pull-down mode" obtains.

When the detector 20 observes a movement frequency 1111111, it is assumed that ordinary video obtains, which must be processed via standard median filtering.

The system, upon detecting a 2:2 or 3:2 "pull-down mode" sets the film mode indication high, which means that the movement detector 11 must be set to insensitive mode and that the processing synchronization must be started. The detection length of the movement detection is also switched over to a shorter word length of, for example, six bits instead of eleven bits.

The abovementioned processing synchronization is necessary to ensure, upon detection of the movement sequences associated with 2:2 or 3:2 "pull-down mode", that processing of the data will proceed in the correct phase.

For the processing, the median filter 1 is switched off and matching fields are merged, and the merged image is repeated, as will be explained below in more detail.

For the purpose of this synchronization, the apparatus includes a synchronizer 21 comprising two oscillating shift registers, one of which, hereinafter referred to as "3:2 pull-down syncer" 22, produces the sequence 10100 and repeats it indefinitely, or in other words continuously produces the 3:2 pull-down processing sequence, while the other one, hereinafter referred to as "2:2 pull-down syncer" 23 produces the sequence 01 and repeats it or, in other words, continuously produces the 2:2 pull-down processing sequence.

The output of the synchronizer 21 connects, via 21A, to the abovementioned memory controller 9.

When one of the abovementioned film modes is detected by the detector 20, the relevant pull-down syncer 22 or 23 is synchronized by a synchronization pulse coming from the film mode/video detection, for example from the detector 20. From that instant, the processing operation is subjected to the activated syncer 22 or 23

which indeed runs entirely synchronously with the incoming film phase.

The activated syncer 22 or 23 also serves as a monitor reference for the film mode/video detection. When the syncer 22 or 23 outputs a zero, and the detector 20 outputs one, an error has crept into the film phase (malfunction in the video source, too many vertical on/off structures, situation of video editing, incorrectly fitted "cue flashes" etc). In such a situation, the apparatus will go into film resync mode, the movement detector 11 being set to very high sensitivity, and the processing will temporarily be switched back to median filtering. For the purpose of this error detection of the film phase, the output of the synchronizer 21 forms part of a feedback loop, via 21B, to the detector 20.

These measures ensure that even if the movement is slight and brief, there is still a good chance of rapidly retrieving the new film phase. If there is too much movement, the movement detector 11 will become saturated, and there is a risk that the apparatus can no longer distinguish between film mode and video mode.

The apparatus will then switch over to normal video mode, i.e. normal median filtering, and will again start to search for film mode using an eleven-bit word length for the history of the movement detector 11 and a nominal sensitivity. From this position, given sufficient movement, the correct film phase will be found again relatively rapidly. The switch-over in the sensitivity of the movement detector 11 can be disabled via an interface. There is then a choice of three fixed settings of the sensitivity.

The film processing proper takes place as follows:

A zero-to-one transition of the movement detector 1 means that movement has been detected between the successive fields. A one-to-one transition means continuous movement between the successive fields. A zero-to-zero transition means standstill between the successive fields. A one-to-zero transition means a transition from movement to standstill between the successive fields.

If the correct film mode has been found and the syncer 22 or 23 is in the correct phase, the requirement, when the output of the syncer 22 or 23 is one, is for merging to take place of the successive fields, whereas, when the output of the syncer 22 or 23 is zero, repetition of the merged field must take place, as schematically indicated in the bottom section of Figure 2 for the 2:2 pull-down or, in the bottom section of Figure 4, for the 3:2 pull-down.

In the case of 2:2 pull-down, the fields A and B, C and D etc. are merged, the line B_0 being introduced between the lines A_0 and A_1, the line B_1 between the lines A_1 and A_2, etc. The merged fields are each time repeated once, as represented by arrows at the bottom in Figure 2.

In the case of 3:2 pull-down, the fields A and A', B and B' etc. are merged, the line A'_0 being introduced between the lines A_0 and A_1, the line A'_1 between the lines A_1 and A_2, etc. The merged fields are each time repeated once or twice, as represented by arrows at the bottom in Figure 4.

Where in normal video mode the result of the median filter 1 is combined with the original lines of the current field, the lines of the current field are now, via 21A and the memory controller 9, combined with the original lines of the next field, so that an image of 625 or 525 lines is produced which is simply the

merging of the even and odd field from one and the same image of the film.

What is introduced between the successive lines is therefore not the median but the line of a following field. In the case of 2:2 pull-down, for example, the line B_0 is introduced between the lines A_0 and A_1, as represented in Figure 2, and in the case of 3:2 pull-down, as represented in Figure 4, the line A'_0 is introduced, for example, between the lines A_0 and A'_1.

In the case of film mode, the median filter 1 is therefore solely used for measuring but not for processing.

This result of the merger is relayed to display bank 2.

The abovementioned information is also relayed, via the median filter 1, to the interpolator 5, so that the interpolation is also performed on the merged fields. The result of the interpolation is relayed to the display bank 6.

The display banks 2 and 6 are read out as described above.

The repetition of the merged fields, which is represented at the bottom in Figures 2 and 4 by arrows, is effected by preventing writing to the display banks 2 and 6, so that, when these display banks 2 and 6 are read out, the same so-called "old" information is read out.

The matching even and odd fields coming from one and the same film image are merged once more until the original film image is obtained, and this image is repeated until it is again possible for a following original film image to be constructed by means of the abovementioned merger.

As already stated previously, the edges and the details of the video images can be enhanced selectively by a so-called "edge boost".

Such an enhancement is performed on the incoming signal by means of two phase-linear "finite impulse response" or FIR filters having a sum of coefficients of 0, the one being a band-pass filter having the coefficients -1 0 2 0 -1 and the other a high-pass filter having the coefficients -1 2 -1. The sample frequency is 16 MHz.

The results of the two filters are merged and scaled, for example divided or multiplied, and added to the original video signal. Both filters, for example, have eight settings. This allows for various combinations of the two filters.

If excessively pronounced edges are detected, the filter effect can be attenuated, if required, to prevent deadlocks.

The invention is by no means limited to the embodiments described hereinabove and represented in the figures, such a method and apparatus for video processing being capable of being implemented in diverse variations without falling outside the scope of the invention.